DOCUMENT RESUME

ED 193 101 SP 016 733

AUTHOR Feldhusen, John F.
TITLE Experimental Research on Teaching and the Evaluation of Teaching in Higher Education.
PUB DATE Apr 80
NOTE 17p.: Paper presented at the American Educational Research Association conference (Boston, MA, April 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classroom Environment; *Componential Analysis; *Educational Research; *Experimental Programs; Higher Education; Instructional Materials; *Instructional Systems; Student Evaluation of Teacher Performance; Student Teacher Relationship; Teacher Behavior; *Teacher Evaluation; *Teaching Methods

ABSTRACT

A review of the concept of complex instructional treatments demonstrates that componential analysis is needed to properly identify the effective elements of an instructional system. Research should focus on both the system and the effectiveness of its parts. Research on teaching also requires that the complex interaction of the teacher's behavior with student behavior, instructional materials, and classroom environment be controlled or assessed. (Author/CJ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED193181

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John F. Feldhusen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

OCT 1980
RECEIVED
ERIC/TM

Experimental Research on Teaching and the Evaluation of Teaching In Higher Education

John F. Feldhusen

Purdue University

I am pleased to have this chance to talk to you today about experimental research on college and university teaching but frightened by the prospect that my audience includes many people who are doing excellent theoretical and empirical work in this area. There are many people here who are far better scholars than I am in this domain of research on college and university teaching. Nevertheless I have made some effort to review literature and my own limited experience and to draw some conclusions about desirable directions for future research. I hope that I will be able to pursue some of them myself in future research activity.

Exactly fifteen years ago I first became aware of the fact that many of our treatments in applied educational research were really extremely complex packages of treatment aspects. This came about during the programmed instruction movement when we mistakenly assumed that the machine and the program could easily be compared with "traditional" modes of teaching. In short order we discovered that both sides of the comparison,

---

AERA Symposium Paper, Boston, April, 1980.

SP 016 733

the programmed instruction and the so called "traditional" mode were really complex and unknown packages of treatment conditions. An article by Lumsdaine in the first Handbook of Research on Teaching (1963) described a technique called componential analysis which specifically addressed this problem, and we soon set off on some research with a creativity training program which we had developed (Feldhusen, Treffinger & Moore, 1970). The results indicated that the parts of our program were indeed differentially effective. We also learned some techniques for analyzing a complex treatment package. In a recent review of research on college teaching, Kulik and Kulik (1979) again remind us of the need to analyze and assess the effects of complex instructional treatments or systems componentially.

The components of instruction may include the type and frequency of tests, the nature and immediacy of feedback, the type and relevance of media including print materials, practice condition, and so forth. The audio-tutorial method, for example, is a very complex treatment package which has typically been researched as a global package while PSI has received a great deal of componential analysis.

Specifying treatment conditions from the instructor-experimenter's point of view may fail to identify the real treatment conditions as they effect students. Crowe and Feldhusen (1976) assessed student's perceptions of course-treatment conditions in order to specify treatments more precisely. We wre painfully aware, through direct experience,

of students who described audio-tutorial as a course where you simple "listen to all those tapes" and PSI as a course in which you "read something and take a test." Students may indeed be failing to perceive and respond to other treatment conditions.

As a further effort to specify treatment-learning conditions we tried in another study (Touq & Feldhusen, 1974) to use student observations and have outside experts view instructor behavior to determine if they agreed on the instructor behaviors. We found agreement quite low.

The instructional treatment should, of course, be viewed as a system as Kozma et al. (1978) remind us, but we still need to know the system as a set of components and how each contributes to the whole system. However, systems analysis also reminds us that we need to know whether the package is really functioning harmoniously as a system. Occasionally we may discover dysfunctional elements such as the total self-pacing element in some PSI systems which leads to severe student procrastination effects.

In summary, componential analysis is needed in all applied educational research on college teaching as an approach to specification of the true and effective treatment conditions.

A major component of any instructional system is the instructor and specifying the nature and the behavior of this element of a treatment may be consummately difficult. I am conducting an experiment in eleven schools at the present time on a method of teaching gifted students in pullout programs.

The treatment package is a complex of materials and methods, all to be implemented by individual teachers. Observations of the teachers reveals that X is implementing our model superbly well while Y simply cannot handle some elements. In the final analysis of this work we will need careful analysis of the teacher component in order to interpret results adequately.

Rosenshine and Furst (1973) reviewed systems for observing teacher behaviors. A number of good techniques are available. From the Tom and Cushman (1975) study we know that different teacher behaviors can make a difference in student learning. Thus, we may indeed discover that a significant part of the treatment effect is attributable to the instructor and not to the treatment package which we think is the heart of the system.

It is natural to turn from specification of the instructor to specification of the subjects and subgroups of subjects. Subgroups may interact differentially with the treatment or its components. It is obvious that students differ, and they differ in their response to different modes of instruction. Research on learning styles and cognitive styles verifies two major dimensions among many. Margraine (1978) reviewed the research in this area and concluded, especially from the work of Entwistle and Brennan (1974) that we should try to identify clusters or subgroups for analysis of the interactions of instructional treatments with student characteristics. Students differ substantially in their motivational patterns and Moen

and Doyle (1978) have recently provided a review of methods of conceptualizing and measuring those motivations.

Applied educational research at the college level may be particularly plagued by student motivation of a goal setting nature which leads to study-learning behaviors which compensate for weaknesses or strengths in instructional treatment conditions. Most students set attainment goals, often as grades, and they probably adjust their study behaviors accordingly. Thus, students in traditional instructional treatments may perform as well as students in a new and theoretically superior treatment because they have set a goal and will do whatever is necessary to earn the grade. The ubiquitous nonsignificant differences may be largely attributable to this effect.

Thus, there is a need to know about out-of-class study time and conditions. Are other people intervening to help in some way? In an experiment on teaching machines at the University of Wisconsin we learned that well meaning TAs, who did not know about the experiment, were tutoring some students in the control group. In one of my field experiments on creativity training in Elkhart, Indiana some control teachers secured experimental teaching material and began using it. In yet another experiment with Purdue students, we asked them to write a narrative description of their study behaviors - after an experiment had ended. Knowing what goes on as study-learning behavior, especially compensating behavior, may be crucial in interpreting results of an experiment. College and university students may differ widely in their approaches to,

and the time they spend on, out-of-class study.

There is also the problem that the students' response to in-class treatments may be essentially passive while the experimenter hopes for active learning responses. Many college students view class time as coming to get information and recording it in notes or on audio tape. Study-learning time occurs later in the library or a dorm room. Thus, we found some students persistently taking notes during a programmed instruction project, rather than giving themselves totally to the learning experience as we wanted. The student who viewed the audio-tutorial method as simply listening to tapes was probably exhibiting this same passivity.

Analysis of students then, as interactors with treatment, is crucial in research on instruction at the college level. Much is known about student characteristics and motivation. This knowledge can be used to design experiments which assess interactions of student variables with treatment conditions. Practically our best bet may be to try to identify student clusters and assess differential impacts on different clusters.

We turn next to the need for analysis of documents or print material in research on teaching. In some instructional treatments the majority of the effect may be expected to come from documents. The program in PI and CAI, the unit documents in PSI, and the minicourses in AT are all powerful carriers of instruction. They differ in quality and design. They may refer students to other print sources such as textbooks and articles which also differ in quality and design. Recent

research on prose learning has shown us that analysis of the documents is manageable and profitable.

Analysis of documents also leads us to concern for the subject mater or curriculum being purveyed in an instructional treatment. Some programmed instruction was developed for defunct curriculum material. We cannot assume that instructional designers are adept at analysis and specification of the curricular content. Ware and Williams' (1977) analysis of "teaching points" in their analysis of Dr. Fox's teaching content was, in my view, a quite inadequate way of specifying the content of an instructional treatment, but to their credit, they did at least try. Other modes of analysis of content such as Gagne's (1970) hierarchical analysis and Markle and Tiemann's (1969) concept analysis along with the Bloom Taxonomy (1965) provide systems for researchers to specify the nature of the curricular content.

Hopefully by being better able to specify the nature of curricular content of an instructional treatment we can approach better generality of the treatment. Gage (1979) has recently discussed this problem of generalization of results of research on teaching as depending very much on grade level and curricular content.

We turn next to the setting or environmental conditions of instructional treatments and how to specify them. In research on college and university teaching, a major variable in this domain, clearly linked to achievement, is class size. The smaller class is quite clearly supportive of discussion

and interactions which facilitate rapport between students and instructors, and it facilitates learning of higher level cognitive skills which require in-class practice.

Student perceptions of the environment of learning are noted by McKeachie and Kulik (1975) as missing elements in reports of research on instruction at the college and university level. Mathis and McGaghie (1972) reported rich descriptions of instructors and the educational settings in which they taught. McKeachie and Kulik (1975) also cite the work of Mann et al. (1970) in their book The College Classroom which provides elaborate description of the interactional setting of four college classrooms. The work of Crowe and Feldhusen (1976) was also cited earlier as illustrative of students' perceptions of the classroom environment.

Methods for the assessment of classroom environments were reviewed by Trent and Cohen (1973). They cite the work of Pace with the College Characteristics Index which documents the environmental presses of the classroom. These include presses for intellectual, independent, interpersonal or group oriented behavior. They also cite the Environmental Assessment technique developed by Astin (1962) which focuses on such variables as class size, average intelligence, realistic orientation, intellectuality, social emphasis, conventionality, artistic orientation, and enterprising qualities. Other approaches are also well developed for assessing - describing the environmental setting of instructional research, but, as McKeachie and Kulik noted (1975), they are rarely used in

instructional research. It seems likely that the setting or environment in which an instructional treatment is embedded may interact with the treatment to effect outcomes. Studies of the interaction of treatments with instructional environments could be of value in clarifying reasons for the effects or lack of effects.

We turn next to the dependent variables in research on instructional treatments at the college and university level. Is the final examination the ultimate criterion? Does a multiple choice hodge podge of knowledge represent significant or worthwhile learning in a course? Are there more desirable alternatives?

Kulik and Kulik (1979) review evidence on time to learn as a significant alternative criterion variable. Long range retention measures shocked us in one of our studies (Speedie, Treffinger & Feldhusen, 1971) when we discovered treatment effects, assessed at the immediate culmination of a treatment, disappearing several months later. Equally important is transfer or application of treatment learning in practical tasks. Rarely are these alternative criteria assessed. In still other studies it might be appropriate to assess attitudinal, social or personal changes as a result of an instructional treatment, but the final exam reigns supreme as the all too frequent criterion measures. However, Kulik and Kulik (1979) note instances of use of integrative essays as criterion measures and other instances of use of measures of higher level thinking can be found.

Closely related to this discussion of the criterion measures is the essential problem of the validity of the criterion measures for the treatment conditions. Criterion measures should assess what the treatment purports to teach. An assessment of this nature can be done using the Bloom Taxonomy (1965) as a vehicle for analysis of treatment goals and test content. Analysis in terms of a hierarchical analysis à la Gagne (1970) or types of learning outcomes such as concepts, principles and problem solving skills would also be desirable. In summary the criterion measures should assess the instructional treatment in terms of all purported goals as well as others which may not have been intended but which may be serendipitous outcomes unanticipated by the designers of the treatment.

An instructional treatment should also be assessed in terms of what we have called instructional validity (Feldhusen, Hynes & Ames, 1976). We assert that instruction has validity only if certain minimal conditions are delineated. These include (1) a specification of the skills or behaviors to be learned, (2) clear evidence that the treatment provides necessary modelling or demonstration of the behaviors to be learned, (3) essential information transmission, (4) opportunity for students to practice behaviors to be learned, (5) and feedback or reinforcement relevant to performance. A study of discussion methods in teaching may give us no clear evidence concerning 1, 2, 3 and 5 as listed above.

We have devoted most of our discussion to research on

instruction. We shall close with comments on several closely related issues. One of these is research on student ratings of instruction. Much more work is needed in this area to clarify remaining issues and problems. Particularly needed are experimental studies in which conditions of ratings are manipulated. We and others have experimentally manipulated the purpose and assessed effects. Others have assessed the effects of teacher expressiveness on ratings (Ware & Williams, 1977). Feldman (1977) reviewed research on reliability of student ratings while Kulik and McKeachie (1975) and Doyle (1975) have reviewed all aspects of research on ratings. Especially necessary is further research on use of ratings in teaching improvement efforts or in influencing instructor behavior. Erickson (1979) reviewed some of this research and reports her own work in this area. Kulik and McKeachie (1975, p. 236) concluded that the work in this area is still a major disappointment. The extensive experimental work of Ware and Williams (1975) and Centra (1979) provides guidelines for further experimental work related to manipulation of instructional conditions and student ratings.

We close by noting yet one more promising aspect of research which should engage attention of experimenters of college teaching. That is the new method of meta-analysis pioneered by Glass (1976) which provides an excellent methodology for assessing all prior research within a domain and thereby deriving better insights for future research. Kulik, Kulik and Cohen's (1979) reviews of research on the

audio-tutorial and the personalized systems of instruction provide the best evaluations to date of the research in these two domains. Meta-analysis also permits generality as well as specificity of analysis in that meta-analysis can be applied to subsets of a group of related studies to assess particular treatment conditions or criterion measures. The meta-analyses of audio-tutorial and PSI research reported by Kulik, Kulik and Cohen (1979a, 1979b) provide valuable insights for future research on instruction at the college and university level.

References

Astin, A. W. An empirical characterization of higher educational institutions. Journal of Educational Psychology, 1962, 53, 224-235.

Bloom, B. S. Taxonomy of educational objectives, Handbook I, Cognitive domain. New York: McKay, 1956.

Centra, J. A. Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions & improvement. San Francisco: Jossey-Bass, 1979.

Crowe, M. H., & Feldhusen, J. F. Student perceptions of organizational features in relationship to course ratings. Contemporary Educational Psychology, 1976, 1, 376-383.

Doyle, K. O. Student evaluation of instruction. Lexington, Mass.: Heath, 1975.

Entwistle, N. S., & Brennan, T. The academic performance of students. British Journal of Educational Psychology, 1971, 41, 268-276.

Erickson, B. L. Improving college teaching. Journal of Higher Education, 1979, 50, 670-683.

Feldhusen, J. F., Hynes, K. P., & Ames, C. A. Is a lack of instructional validity contributing to the decline of achievement test scores? Educational Technology, 1976, 16, 13-16.

Feldhusen, J. F., Treffinger, D. J., & Moore, S. J. Global and componential evaluation of creativity instructional materials. Buffalo: Creative Educational Foundation, 1970.

Feldman, K. A. Consistency and variability among college students in rating their teachers and courses. Research in Higher Education, 197 , 6, 223-274.

Gage, N. L. The generality of dimensions of teaching. In P. L. Peterson and H. J. Walberg (Eds.), Research on teaching. Berkeley: McCutcheon, 1979.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1970.

Glass, G. V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 10, 3-8.

Kozma, R. B., Belle, L. W., & Williams, G. W. Instructional techniques in higher education. Englewood Cliffs, NJ: Educational Technology Publications, 1978.

Kulik, J. A., & Kulik, C. C. College teaching. In P. L. Peterson & H. J. Walberg (Eds.), Research on teaching. Berkeley: McCutcheon, 1979, 70-93.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. Research on audio-tutorial instruction: A meta-analysis of comparative studies. Research In Higher Education, 1979a, 11, 321-341.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. A meta-analysis of outcome studies of Keller's Personalized System of Instruction. American Psychologist, 1979b, 34, 307-318.

Lumsdaine, A. A. Instruments and media. In N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963, 583-682.

Mann, R. D., Arnold, S. M., Buider, J., Cytrynbaum, S., Newman, B. M., Ringwald, B., Ringwald, S., and Rosenwein, R. The college classroom: Conflict, change and learning. New York: Wiley, 1970.

Margraine, S. A. Student characteristics and academic performance in higher education. Research In Higher Education, 1978, 8, 111-123.

Markle, S. M., & Tiemann, P. W. Really understanding concepts. Champaign, IL: Stipes, 1969.

Mathis, B. C., & McGaghie, W. C. Profiles in college teaching. Evanston, IL: Center for the Teaching Professions, 1972.

McKeachie, W. J., & Kulik, J. A. Effective college teaching. In F. N. Kerlinger (Ed.), Review of research in education 3. Itasca, IL: Peacock, 1975.

Moen, R., & Doyle, K. O. Measures of acedemic motivation. Research In Higher Education, 1978, 8, 1-23.

Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R. M. W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1974, 122-183.

Speedie, S. M., Treffinger, D. J., & Feldhusen, J. F. Evaluation of components of the Purdue Creative Thinking Program: A longitudinal study. Psychological Reports, 1971, 29, 395-398.

Tom, F. K., & Cushman, H. R. The Cornell diagnostic observation and reporting system for student description of college teaching. Search, 1975, 5(8), 1-25.

Touq, M. S., & Feldhusen, J. F. Validity of student ratings of instructors. College Student Journal, 1974, 8, 2-5.

Trent, J. W., & Cohen, A. M. Research on teaching in higher education. In R. M. W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973, 997-1071.

Ware, J. E., Jr., & Williams, R. G. Disciminant analysis of student ratings as a means for identifying lecturers who differ in enthusiasm or information-giving. Educational and Psychological Measurement, 1977, 37, 627-639.